Synthetic MRI of the Knee: Phantom Validation and Comparison with Conventional MRI

Kumar, Neil M ; Fritz, Benjamin ; Stern, Steven E ; Warntjes, J B Marcel ; Lisa Chuah, Yen Mei ; Fritz, Jan

Abstract: Purpose To test the hypothesis that synthetic MRI of the knee generates accurate and repeatable quantitative maps and produces morphologic MR images with similar quality and detection rates of structural abnormalities than does conventional MRI. Materials and Methods Data were collected prospectively between January 2017 and April 2018 and were retrospectively analyzed. An International Society for Magnetic Resonance in Medicine-National Institute of Standards and Technology phantom was used to determine the accuracy of T1, T2, and proton density (PD) quantification. Statistical models were applied for correction. Fifty-four participants (24 men, 30 women; mean age, 40 years; range, 18-62 years) underwent synthetic and conventional 3-T MRI twice on the same day. Fifteen of 54 participants (28%) repeated the protocol within 9 days. The intra- and interday agreements of quantitative cartilage measurements were assessed. Contrast-to-noise (CNR) ratios, image quality, and structural abnormalities were assessed on corresponding synthetic and conventional images. Statistical analyses included the Wilcoxon test, test, and Cohen Kappa. P values less than or equal to .01 were considered to indicate a statistically significant difference. Results Synthetic MRI quantification of T1, T2, and PD values had an overall model-corrected error margin of 0.8%. The synthetic MRI interday repeatability of articular cartilage quantification had native and model-corrected error margins of 3.3% and 3.5%, respectively. The cartilage-to-fluid CNR and menisci-to-fluid CNR were higher on synthetic than conventional MR images (P .001, respectively). Synthetic MRI improved short-tau inversion recovery fat suppression (P .01). Intermethod agreements of structural abnormalities were good (kappa, 0.621-0.739). Conclusion Synthetic MRI of the knee is accurate for T1, T2, and proton density quantification, and simultaneously generated morphologic MR images have detection rates of structural abnormalities similar to those of conventional MR images, with similar acquisition time.

DOI: https://doi.org/10.1148/radiol.2018173007

Synthetic MRI of the Knee: Phantom Validation and Comparison with Conventional MRI

Neil M. Kumar, MD • Benjamin Fritz, MD • Steven E. Stern, PhD • J. B. Marcel Warntjes, PhD • Yen Mei Lisa Chuah, PhD • Jan Fritz, MD

From the Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins University School of Medicine, 601 N Caroline St, Baltimore, MD 21287 (N.M.K., J.F.); Department of Radiology, Balgrist University Hospital, Zurich, Switzerland (B.F.); Faculty of Medicine, University of Zurich, Zurich, Switzerland (B.F.); Bond Business School, Bond University, Gold Coast, Australia (S.E.S.); Center for Medical Imaging Science and Visualization, Linköping University, Linköping, Sweden (J.B.M.W.); Division of Clinical Physiology, Department of Medicine and Health, University Hospital, Linköping, Sweden (J.B.M.W.); SyntheticMR AB, Linköping, Sweden (J.B.M.W.); and Siemens Healthcare GmbH, Erlangen, Germany (Y.M.L.C.). Received December 20, 2017; revision requested February 19, 2018; revision received July 12; accepted July 16. **Address correspondence to** J.F. (e-mail: *jfritz9@jhmi.edu*).

SyntheticMR AB and Siemens Healthcare provided intellectual and technological support. This project was performed at the Johns Hopkins Medical Institutions, in collaboration with SyntheticMR AB and Siemens Healthcare under a master research agreement, and in accordance with the Johns Hopkins standards for industrial partnership. Johns Hopkins physician researchers, including myself, had full control over the data at any point in time and guarantee the accuracy of the presented data and integrity of this study.

Conflicts of interest are listed at the end of this article.

Radiology 2018; 00:1–13 • https://doi.org/10.1148/radiol.2018173007 • Content code: MK

Purpose: To test the hypothesis that synthetic MRI of the knee generates accurate and repeatable quantitative maps and produces morphologic MR images with similar quality and detection rates of structural abnormalities than does conventional MRI.

Materials and Methods: Data were collected prospectively between January 2017 and April 2018 and were retrospectively analyzed. An International Society for Magnetic Resonance in Medicine−National Institute of Standards and Technology phantom was used to determine the accuracy of T1, T2, and proton density (PD) quantification. Statistical models were applied for correction. Fifty-four participants (24 men, 30 women; mean age, 40 years; range, 18–62 years) underwent synthetic and conventional 3-T MRI twice on the same day. Fifteen of 54 participants (28%) repeated the protocol within 9 days. The intra- and interday agreements of quantitative cartilage measurements were assessed. Contrast-to-noise (CNR) ratios, image quality, and structural abnormalities were assessed on corresponding synthetic and conventional images. Statistical analyses included the Wilcoxon test, χ^2 test, and Cohen Kappa. *P* values less than or equal to .01 were considered to indicate a statistically significant difference.

Results: Synthetic MRI quantification of T1, T2, and PD values had an overall model-corrected error margin of 0.8%. The synthetic MRI interday repeatability of articular cartilage quantification had native and model-corrected error margins of 3.3% and 3.5%, respectively. The cartilage-to-fluid CNR and menisci-to-fluid CNR were higher on synthetic than conventional MR images ($P \leq .001$, respectively). Synthetic MRI improved short-tau inversion recovery fat suppression ($P < .01$). Intermethod agreements of structural abnormalities were good (kappa, 0.621–0.739).

Conclusion: Synthetic MRI of the knee is accurate for T1, T2, and proton density quantification, and simultaneously generated morphologic MR images have detection rates of structural abnormalities similar to those of conventional MR images, with similar acquisition time.

MRI signal is described by hardware-specific factors, proton density (PD) signal scaling factors, voxel volume, and pulse-sequence weighting, whereas quantitative MRI uses Bloch-Torrey equations that govern pulse sequence weighting components to distill tissue-specific properties such as T1, T2, and PD (1,2). Based on quantitative MRI data, synthetic MRI can generate both qualitative and quantitative images simultaneously from parental data (2–8).

Quantitative MRI mapping techniques of the knee allow the early characterization and quantification of articular abnormalities and effects of therapeutic interventions (9,10), whereas morphologic T1-weighted, intermediate-weighted, T2-weighted, and short-tau inversion recovery (STIR) MR images allow the characterization of structural abnormalities. However, the separate acquisitions of quantitative and qualitative images can be time consuming, and thus synthetic MRI may be advantageous by offering the simultaneous generation of quantitative maps and morphologic images. Synthetic MRI has been successfully used in the brain (11,12), but its role is less well established for MRI of the knee.

We tested the hypothesis that synthetic MRI of the knee generates accurate and repeatable quantitative maps and produces morphologic MR images with similar detection rates of structural abnormalities as conventional MRI.

Materials and Methods

Employees of SyntheticMR AB (Linköping, Sweden) (J.B.M.W.) and Siemens Healthcare (Erlangen, Germany) (Y.M.L.C.) provided intellectual and technological support. Authors (N.M.K., B.F., S.E.S., J.F.) who were not employees of or consultants for SyntheticMR AB and Siemens Healthcare performed the data evaluations and had control of inclusion of any data and information that

Abbreviations

CNR = contrast-to-noise ratio, ISMRM = International Society for Magnetic Resonance in Medicine, NIST = National Institute of Standards and Technology, PACS = picture archiving and communication system, PD = proton density, ROI = region of interest, SNR = signal-to-noise ratio, STIR = short-tau inversion recovery

Summary

Synthetic QRAPMASTER MRI of the knee is accurate for T1, T2, and proton density quantification, and simultaneously generated synthetic morphologic MR images have detection rates of structural abnormalities similar to those of conventional MR images, with similar acquisition time.

Implications for Patient Care

- Based on quantitative QRAPMASTER data, synthetic MRI of the knee generates quantitative maps and morphologic MR images with the same acquisition time required for conventional morphologic MRI.
- Synthetic MRI of the knee is accurate for T1, T2, and proton density quantification with phantom-based model-corrected average error margin of 0.8%.
- Synthetically generated morphologic MR images using the QRAPMASTER technique have higher contrast resolution of cartilage and meniscus relative to joint fluid when compared with conventional MRI, and similar detection rates for structural abnormalities as conventional MRI with similar acquisition time.

might have presented a conflict of interest for those authors (Y.M.L.C. and J.B.M.W.) who were employees of or consultants for SyntheticMR AB and Siemens Healthcare.

Our study was approved by our institutional review board and complied both with the Declaration of Helsinki and the Health Insurance Portability and Accountability Act. Written informed consent was obtained from all participants for prospective data collection and retrospective analysis.

MRI Technique

We used a commercially available wide-bore 3-T MRI system (Magnetom Skyra, Numaris/4 Syngo MR E11C; Siemens Healthcare) with 48 independent radiofrequency receiver channels, maximum gradient field amplitude of 45 mT/m, and a slew rate of 200 T/m/sec. For phantom measurements, a head coil (Siemens Healthcare) with 16 receiver channels was used. For human participants, a knee coil (Quality Electrodynamics, Mayfield, Ohio) with one transmit and 15 receiver channels was used.

For the acquisition of parental synthetic MRI data, we used a sagittally oriented, biphasic QRAPMASTER (quantification of relaxation times and proton density by multiecho acquisition of a saturation-recovery using turbo spin-echo readout) MRI pulse sequence prototype (6,7). The pulse sequence used a two-dimensional, multisection, multiecho, multisaturation delay saturation-recovery turbo spin-echo technique with a repetition time of 4000 msec, two echo times of 21 and 103 msec, inversion time of 27 msec, four saturation delay times of 150, 580, 1860, and 3860 msec, parallel acceleration factor of 3, echo train length of five, receiver bandwidth of 401 Hz per pixel, flip angle of 150°, field of view of 160 × 160 mm^2, matrix of 320 × 240, section thickness of 3 mm and 0.3 mm intersection gap, and 28 sections (Table 1). The acquired data were used to generate quantitative T1, T2, and PD maps and qualitative T1-weighted, intermediate-weighted, T2-weighted, and STIR MR images by using commercially and publicly available software (SyMRI NEURO, version 8.0.4; SyntheticMR AB). The commercially and publicly available SyMRI NEURO software package was characterized by advanced functions, including the ability to export and transfer synthesized images in Digital Imaging and Communications in Medicine format to our picture archiving and communication system (PACS) (Vue version 12.1.0.2041; Carestream Health, Rochester, NY) for observer evaluations. Additionally, this software package afforded full synthetic functionality for synthesizing the entire spectrum of quantitative and morphologic musculoskeletal MR images and contrasts, without any restrictions to neuroradiological MRI.

For participants, we additionally acquired conventional T1-weighted, intermediate-weighted, T2-weighted, and STIR MR images with similar parameter settings (Table 1). The total acquisition times for conventional MRI and synthetic MRI for participants were 9 minutes 21 seconds and 9 minutes 52 seconds, respectively.

Phantom Evaluation

To validate the accuracy of the quantitative knee pulse sequence, we used an MRI system phantom developed by the International Society for Magnetic Resonance in Medicine (ISMRM) Ad Hoc Committee on Standards for Quantitative Magnetic Resonance and the National Institute of Standards and Technology (NIST) (13). The ISMRM-NIST phantom was considered the standard of reference, and synthetic MRI was considered the index test. The phantom consisted of T1, T2, and PD layers. Each layer contained 14 spheres with previously determined absolute T1 and T2 and PD percentage values at room temperature. We used spheres 1–6 in the T1 layer (351.5–1989 msec), spheres 1–10 in the T2 layer (22.56–581.3 msec), and spheres 1–14 in the PD layer (5%–100%). The MRI suite was set to 20°C. The bore fan was set on lowest convection. The phantom was given 12 hours to adapt to room temperature.

The ISMRM-NIST MRI phantom data acquisition was performed on 2 consecutive days to assess interday repeatability. On each day, two sessions were performed to assess intraday variability. During each session, each of the three layers was imaged twice at 30-minute intervals to assess intrasession repeatability. In total, each layer of the phantom was imaged eight times. After each session, we repositioned the phantom in the coil and the coil in the MRI system. One observer (J.F.) with 15 years of musculoskeletal MRI experience performed measurements (SyMRI NEURO, version 8.0.4) of T1, T2, and PD values on synthetic T1, T2, and PD maps (Fig 1) using 1 cm^2 round regions of interest (ROIs). All measurements were repeated three times at 1-week intervals.

While the accuracy of the QRAPMASTER technique was assessed with the phantom measurements that were based on Bloch equations, heteroscedastic variation and residual errors of the quantitative data were then addressed through model-based correction by using logarithmic transformation and quadratic and split (segmented) quadratic equations. For PD data, which were expressed as percentage values and demonstrated no heteroscedasticity, logarithmic transformation was not required.

Table 1: MRI Study Protocol

	Synthetic MRI, QRAPMASTER				Conventional MRI, Turbo Spin Echo			
Parameter	T1 Weighted	IW	T2 Weighted	STIR	T1 Weighted	IW	T2 Weighted	STIR
Repetition time (msec)	466	4000	4000	5860	466	4000	4000	5860
Echo time (msec)	7.9	31	102	30	7.9	31	102	30
Inversion time (msec)	—	—	—	220	—	—	—	220
Acceleration	3	3	3	3	1	1	1	1
Echo train length	5	5	5	5	3	15	15	17
Flip angle (degrees)	150	150	150	150	150	150	150	150
Receiver bandwidth (Hertz/pixel)	401	401	401	401	504	466	466	504
Field of view (mm)	160 × 160	160 × 160	160 × 160	160 × 160	160 × 160	160 × 160	160 × 160	160 × 160
Matrix	320 × 240	320 × 240	320 × 240	320 × 240	320 × 240	320 × 240	320 × 240	320 × 240
Section thickness/gap (mm)	3/0.3	3/0.3	3/0.3	3/0.3	3/0.3	3/0.3	3/0.3	3/0.3
No. of signals acquired	1	1	1	1	1	1	1	1
Concatenation	1	1	1	1	2	1	1	1
No. of sections	28	28	28	28	28	28	28	28
Phase-encoding direction	Anterior-to-posterior	Anterior-to-posterior	Anterior-to-posterior	Anterior-to-posterior	Anterior-to-posterior	Anterior-to-posterior	Anterior-to-posterior	Anterior-to-posterior
Acquisition time	…*	…*	…*	…*	1 min 54 sec	2 min 14 sec	2 min 16 sec	2 min 57 sec

Note.—IW = intermediate weighted, STIR = short-tau inversion recovery.

* For synthetic MRI (QRAPMASTER), the combined acquisition time for T1 weighted, IW, T2 weighted, and STIR images was 9 minutes 52 seconds.

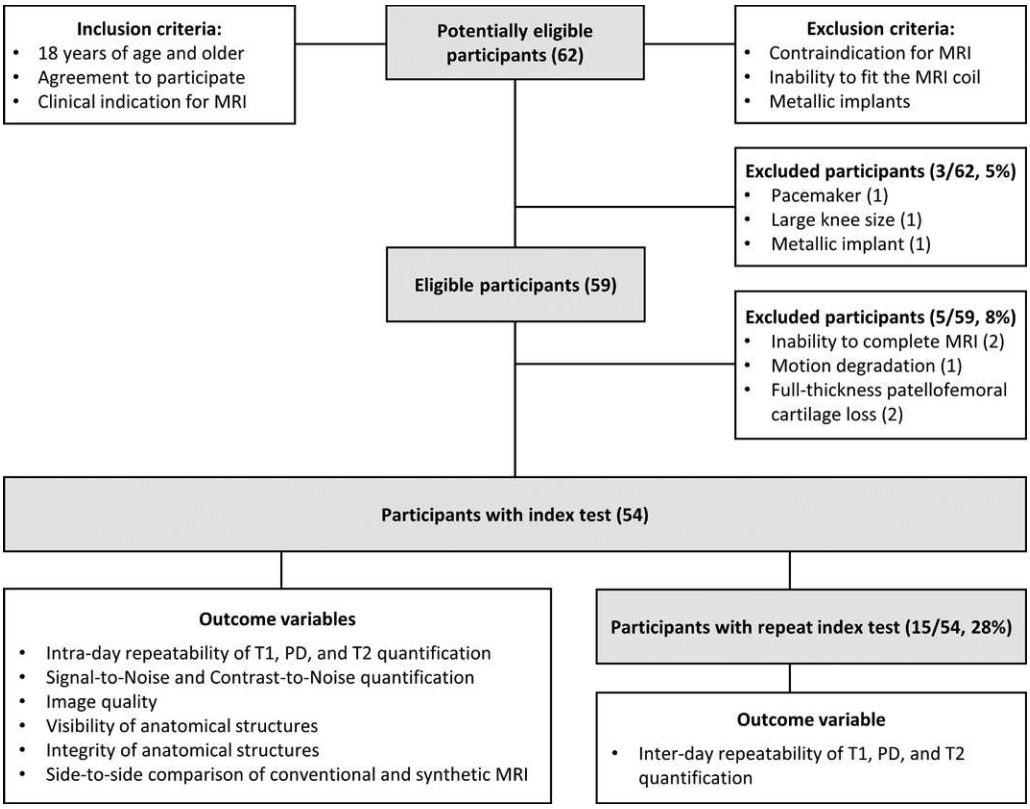


Figure 1: Flow diagram of participants through the study. PD = proton density.

Model-based correction was performed to reduce inhomogeneity of errors across parameter domains to maintain accuracy at the extremes of the included relaxation times. Heteroscedastic variation can occur due to additive Gaussian noise at longer repetition and echo times and monoexponential fitting not accounting for the effects of unmodeled variables such as spatially varying gradients, magnetization transfer, and anisotropy (8,14,15). Models were fit by using residual errors determined by ordinary least-squares, as the transformations induced reasonable homoscedasticity.

Participant Evaluation

Between January 2017 and April 2018, 54 participants (mean age, 40 years; age range, 18–62 years) including 24 men (mean age, 37 years; range, 18–62 years) and 30 women (mean age, 40 years; range, 21–60 years) were recruited from our practice (Fig 1). Indications for knee MRI were made in accordance with published guidelines (16). Each participant underwent our MRI study protocol twice on the same day (Table 1). Between the two acquisitions, participants rested for 30 minutes in a chair. After each acquisition, we repositioned the coil in the MRI system. Fifteen of 54 participants (28%) underwent the MRI protocol again after 5 days on average, with a range of 1 to 9 days. For this study part, conventional MRI was considered the standard of reference and synthetic MRI, the index test. Following data acquisition, the synthetic MRI data were exported to a network drive, imported into the dedicated software (SyMRI NEURO, version 8.0.4) where the quantitative maps and morphologic MR images were synthesized with a semiautomatic preset protocol on a standard desktop computer, and finally sent to our PACS (Vue version 12.1.0.2041; Carestream Health). This process required approximately 5 minutes or less.

Quantitative outcome variables included intraday and interday repeatability of T1, T2, and PD measurements of cartilage on quantitative maps, as well as signal-to-noise (SNR) and contrast-to-noise (CNR) ratios of fluid, cartilage, meniscus, marrow, and muscle on morphologic T1, intermediate-weighted, T2, and STIR images of synthetic and conventional MRI.

T1, T2, and PD value measurements were performed in central patella articular cartilage by one observer (J.F.) (17). When patellar cartilage thickness was insufficient, central trochlear cartilage was measured. The mean pixel value of oval 0.1 cm^2 ROIs sampling approximately 40 pixels was used (SyMRI NEURO, version 8.0.4;). Measurements were repeated three times with 1-day intervals in between measurements.

SNR and CNR were measured (SyMRI NEURO, version 8.0.4) in cancellous bone (distal femoral metaphysis), articular cartilage (patella or alternatively trochlear cartilage), joint fluid (intercondylar notch), and meniscus (posterior horn of the medial or lateral meniscus) by one observer (J.F.). Round or oval ROIs were copied into identical locations on synthetic and conventional MR images. The mean pixel value of the ROIs was used as the signal intensity, whereas the mean standard deviation of a background ROI placed just anterior to the skin surface over the patella was used as the noise. SNR was determined as signal intensity of tissue divided by standard deviation of tissue. Subsequently, CNR was calculated as $|SNR_{(tissue\ 1)} - SNR_{(tissue\ 2)}|$. Measurements were repeated three times at 1-day intervals.

All qualitative outcome variables were obtained by two fellowship-trained musculoskeletal radiologists (B.F. and N.K.), with 5 and 10 years of musculoskeletal MRI experience, respectively. Evaluations were performed independently on randomized data sets after removal of all clinical and personal information. Disagreements of structural integrity and side-to-side comparison assessments were resolved through a final consensus interpretation. Assessments were performed with a standardized, equidistant, five-point Likert scale, where a rating of 1 denoted

"very bad" with complete obscuration of anatomic details, and a rating of 5 denoted "very good" with the unimpaired depiction of all anatomic details. Assessments were performed on PACS software (Vue version 12.1.0.2041; Carestream Health). A 4 × 2 view-port setup was used with synchronized scrolling, sizing, and panning.

Image quality assessments included the degree of motion, noise, artifacts including chemical shift, interface and reconstruction artifacts, edge sharpness of structures, partial volume effects, contrast resolution defined as visual gray-scale differences between structures, fluid brightness, and fat suppression. An interface artifact manifests as an artifactual linear signal along interfaces of different tissues and may occur if there is motion during acquisition.

Visibility of menisci, articular cartilage, cruciate ligaments, extensor tendons, and bone was evaluated in the context of internal derangement assessment on synthetic and conventional data sets consisting of T1-weighted, intermediate-weighted, T2-weighted, and STIR images.

The integrity of menisci, articular cartilage, anterior cruciate ligament, and subchondral bone was assessed. Meniscal tears were defined as substance defect extending to the articular surface. Articular cartilage defects were defined as substance loss greater than 50%. Only the largest articular cartilage defect was assessed. Anterior cruciate ligament tears were defined as 50% or greater substance loss of cross-sectional area. Bone marrow edema was defined as STIR signal hyperintensity compared with the distant normal marrow. Discrepant findings were resolved during consensus interpretation.

Both observers performed a side-to-side comparison, rating corresponding synthetic and conventional T1-weighted, intermediate-weighted, T2-weighted, and STIR images as superior, inferior, or equal based on their subjective impression of the suitability of the images for accomplishing an evaluation for internal knee derangement.

Statistical and Quantitative Assessment

Statistical analyses were performed by using R 3.3 software with lme4 and epiR packages *(http://cran.r-project.org/)*. Variables are given as the average with standard deviation, median with minimum and maximum in parentheses, ratios, or percentages. For the evaluation of the qualitative outcome variables, an apriori Wilcoxon signed-rank test for related samples power calculation derived a minimum sample size of 26 participants for an effect size of 1 for binary or Likert scales, a Bonferroni-corrected alpha error probability of .001, and power of 0.90. Skewness was assessed with the Shapiro-Wilk test. Differences in the comparison assessments in participants were assessed with the Wilcoxon test for related samples or χ^2 test. The coefficient of variation was used to assess the precision of measurements. The interobserver and intermethod agreements in participants were determined by using the Cohen kappa test with linear weights for Likert scale assessments and the Cohen kappa test without weights for binary assessments. Kappa values were graded according to Landis and Koch (18). In the case of acceptable agreement, observer assessments were combined. *P* values less than or equal to .01 were considered to indicate a statistically significant difference.

Results

Phantom Evaluation

The native relative errors of measured T1, T2, and PD values compared with phantom reference values, adjusted for day, session, and replicate variations, were 1.9%, 7.4%, and 5.1%, respectively, whereas the model-corrected relative errors of measured T1, T2, and PD values were 0.8%, 1.4%, and 0.3%, respectively. The average relative error of measured values to reference values was 0.5% (0.1%–0.9%) for intrasession measurements, 0.8% (0.6%–1.2%) for intersession measurements, and 1.0% (0.7%–1.1%) for interday measurements. The average overall relative error of measured values compared with reference phantom values was 0.8% (0.3%–1.4%) following model correction.

The T1 measurements demonstrated a heteroscedastic variation. Fitting of log-linear and basic quadratic nonlinear models yielded the following calibration equation with an accuracy of 0.8% (Fig 2):

$$\widehat{\text{Reference}}\,[\text{msec}] = -36.09 + \sqrt{1302.42 + 1.59 \times \text{Measurement}\,[\text{msec}]^{1.95}}$$

The T2 measurements demonstrated a heteroscedastic variation as well. Fitting of a split quadratic model with a B-splines approach yielded the following calibration equation with an accuracy of 1.4% (Fig 2):

$$\widehat{\text{Reference}}\,[\text{msec}] = \begin{cases} -43.53 + \sqrt{1895.19 + 11.10 \times \text{Measurement}\,[\text{msec}]^{1.62}} \\ \qquad \text{if Measurement} < 195.63 \text{ msec} \\ -935.03 + \sqrt{1020698 + 51.72 \times \text{Measurement}\,[\text{msec}]^{1.62}} \\ \qquad \text{otherwise} \end{cases}$$

The PD measurements demonstrated homoscedastic variation. Fitting of a split quadratic model with a B-splines approach yielded the following calibration equation with an accuracy of 0.3% (Fig 2):

$$\widehat{\text{Reference}}\,[\text{pu}] = \begin{cases} 97.13 - \sqrt{9433.65 - 195.82 \times \text{Measurement}\,[\text{pu}]^{1.01}} \\ \qquad \text{if Measurement} < 32.95 \text{ pu} \\ -24.92 + \sqrt{-1504.66 + 145.95 \times \text{Measurement}\,[\text{pu}]^{1.01}} \\ \qquad \text{otherwise} \\ 164.02 - \sqrt{19994.57 - 155.14 \times \text{Measurement}\,[\text{pu}]^{1.01}} \\ \qquad \text{if Measurement} > 66.30 \text{ pu} \end{cases}$$

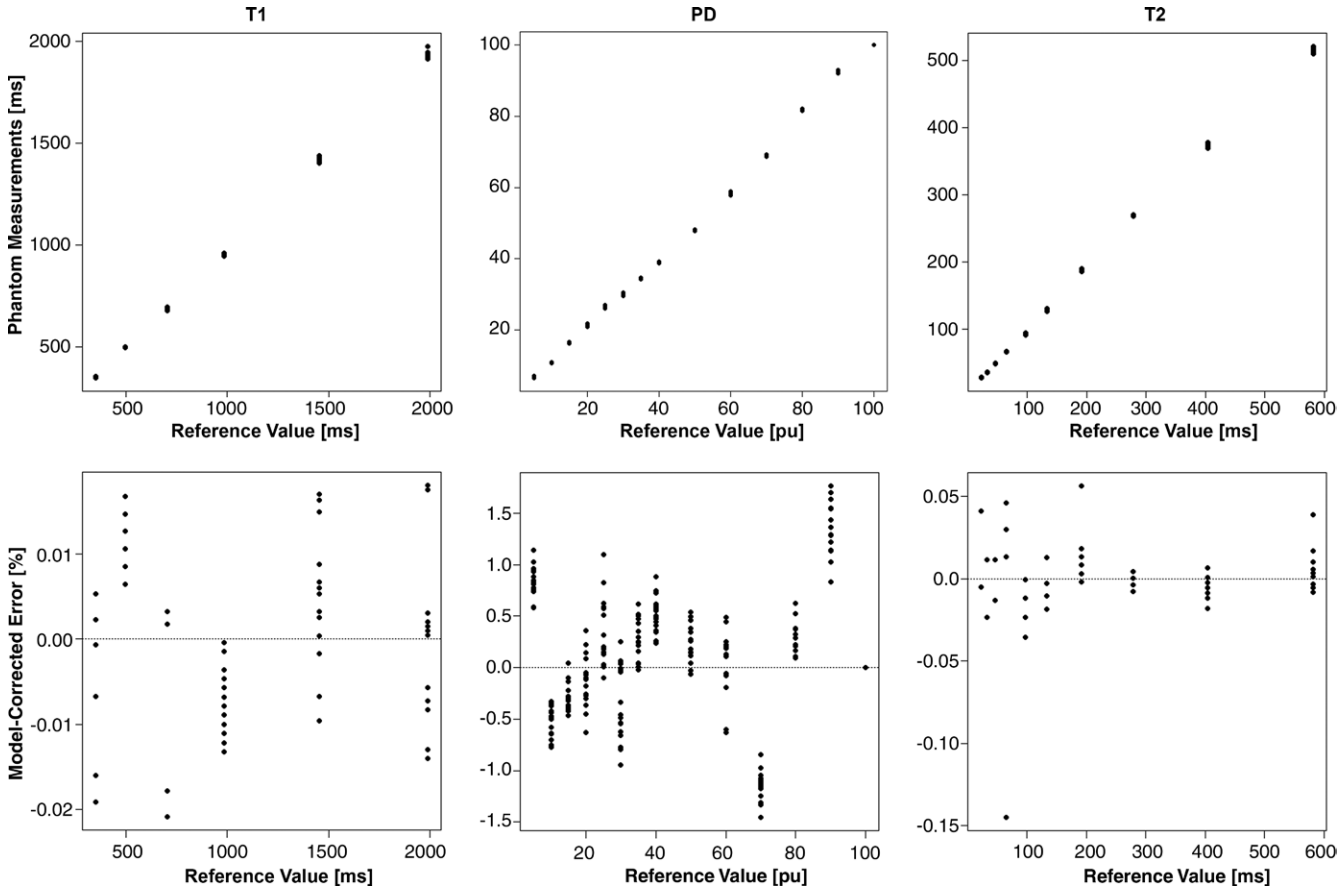


Figure 2: Graphs depict measurement accuracy of synthetic quantitative T1, T2, and proton density (PD) images with an International Society for Magnetic Resonance in Medicine—National Institute of Standards and Technology MRI phantom.

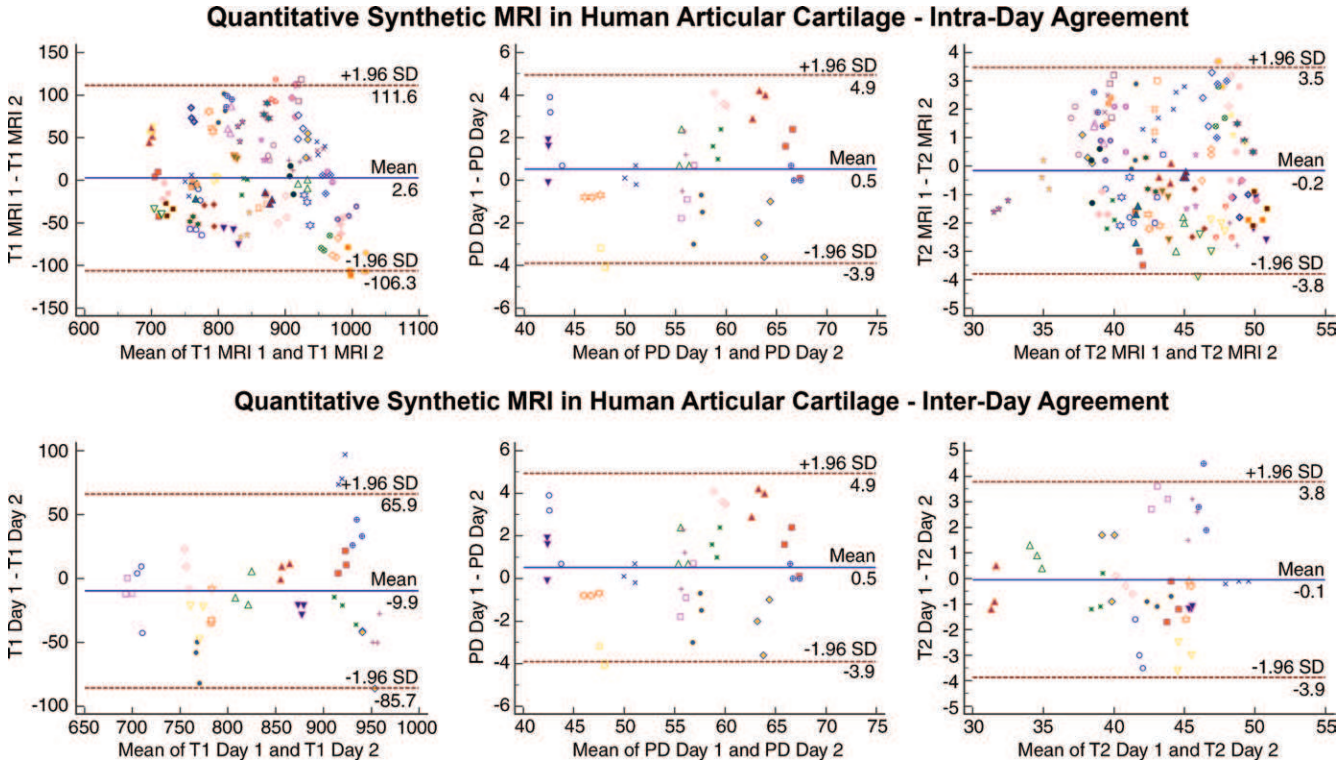


Figure 3: Bland-Altman plots of in vivo intraday and interday agreement of quantitative T1, T2, and proton density (PD) measurements on synthetic MRI quantitative maps. Each symbol represents one participant. Three symbols of the same kind represent three repeat measurements.

Participant Evaluation

Intraday comparison of quantitative T1, T2, and PD measurements of articular cartilage showed an average difference of 4.1% (minimum, 0.1%; maximum, 12.4%) (Fig 3). The coefficient of variation of measurements was 1.1% for both the first and second session.

Interday comparison of quantitative T1, T2, and PD measurements of articular cartilage showed an average difference of 3.3% (0.3%–9.4%) (Fig 3). The coefficient of variation of measurements was 1.2% for both the 1st and 2nd days.

After model correction with phantom data–derived equations, the interday comparison of quantitative T1, T2, and PD measurements of articular cartilage showed an average difference of 3.5% (0.3%–9.6%) (Table 2). The coefficient of variation of model-corrected measurements was 1.3% for both the 1st and 2nd days. The average repeatability coefficient was 21.86 (6.8%).

SNR and CNR ratios of different tissues of morphologic synthetic and conventional T1-weighted, intermediated-weighted, T2-weighted, and STIR MR images and their comparison are given in Figure 4. On synthetic T1-weighted images, SNR of fluid was lower ($P < .001$). On synthetic intermediated-weighted and T2-weighted MR images, SNR of cartilage and SNR of fluid were higher ($P < .001$, respectively). On synthetic STIR images, SNR of fluid was higher ($P < .001$) and SNR of bone marrow and SNR of menisci were lower ($P < .001$, respectively). On synthetic T1-weighted images, the fluid-to-menisci CNR was lower ($P < .001$) and cartilage-to-fluid CNR was higher ($P < .001$). On synthetic intermediated-weighted, T2-weighted, and STIR images, the cartilage-to-fluid CNR, menisci-to-fluid CNR, and muscle-to-fluid CNR were higher ($P < .001$).

Image quality assessments (Table 3) showed synthetic MRI had greater STIR fat suppression ($P < .001$) and fluid signal ($P = .10$), as well as higher degrees of image noise ($P = .001$) and artifacts ($P < .001$) (Fig 5). There were no differences between the other image quality parameters (Table 3).

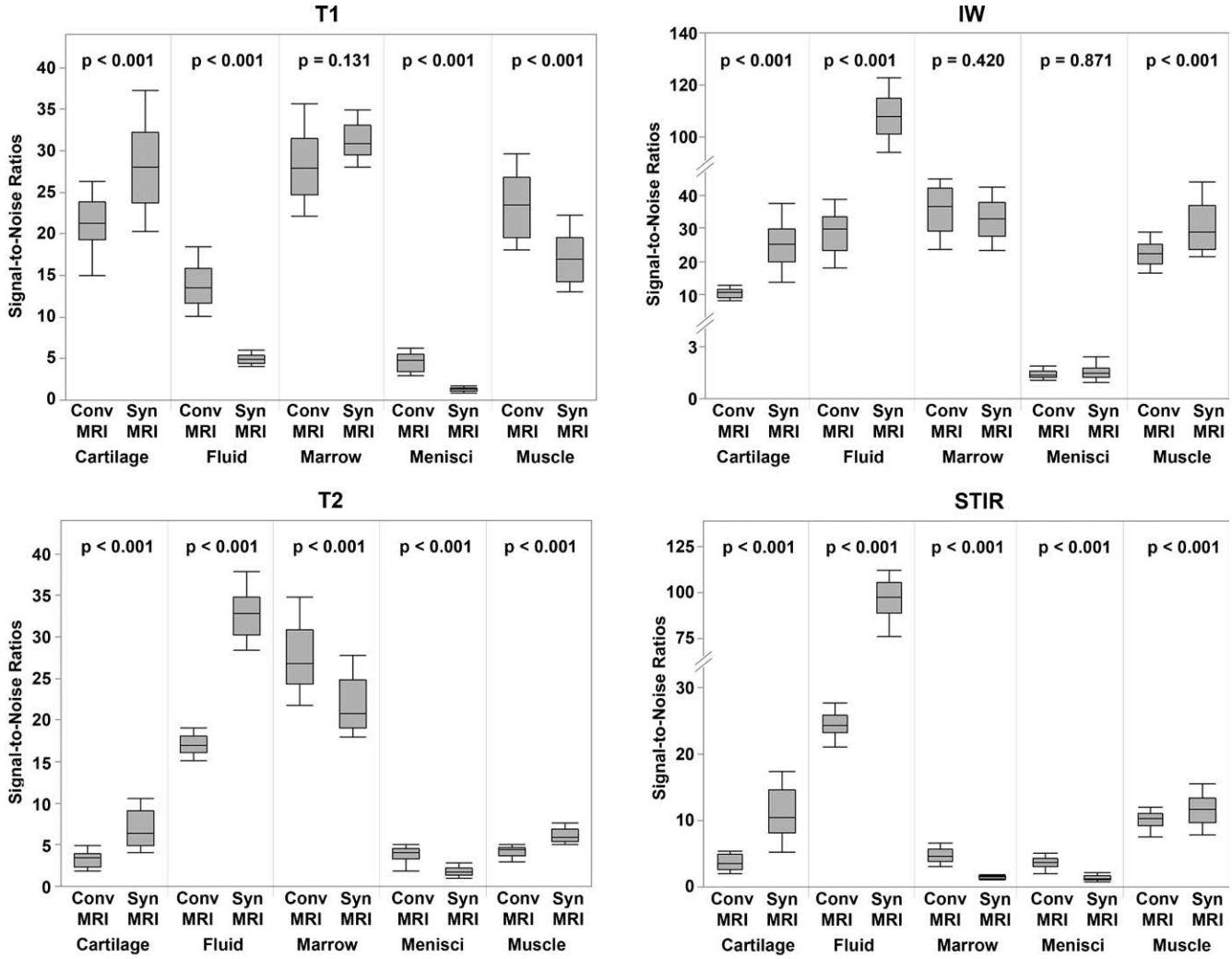
Visibility of menisci, articular cartilage, anterior and posterior cruciate ligaments, extensor tendons, and bone was rated as good to very good on conventional and synthetic STIR, T1-, intermediate-, and T2-weighted MR images, with interobserver agreements ranging from moderate to good (kappa, 0.584–0.708) and no differences noted (P values = .01–.73). Table 4 shows the frequencies of meniscal tears (Fig 6), articular cartilage defects (Fig 5), and areas of bone marrow edema. There were no anterior cruciate ligament and extensor mechanism tears. The interobserver agreements were moderate to very good. The intermethod agreements were good. Among 108 potential discrepancies between conventional and synthetic MRI for both observers of each structure, there were 11 (10%) for medial meniscus, nine (8%) for lateral meniscus, 22 (20%) for articular cartilage defects, and 11 (10%) for bone marrow edema.

For side-to-side comparison, observer A rated synthetic MRI in six of 54 (11%) and conventional MRI in three of 54 (6%) participants as superior, whereas 45 of 54 (83%) were rated as equivalent. Observer B rated synthetic MRI in three of 54 (6%) and conventional MRI in six of 54 (11%) participants as superior, whereas 45 of 54 (83%) were rated as equivalent ($\chi^2 = 16$, $P = .003$).

Table 2: In Vivo Repeatability of Raw and Model-corrected Quantitative Synthetic MRI Data

	Native Synthetic MRI Data					Model-corrected Synthetic MRI Data					
Parameter	Average Difference (%)*	Minimum Difference (%)	Maximum Difference (%)	CV Day 1 (%)	CV Day 2 (%)	Average Difference (%)*	Minimum Difference (%)	Maximum Difference (%)	CV Day 1 (%)	CV Day 2 (%)	*P* Value of Native and Model-corrected Data Comparison
T1	3.4 ± 2.6	0.8	9.4	1.0	1.0	3.6 ± 3.2	0.4	10.1	1.0	1.1	.27
PD	3.2 ± 2.3	0.3	7.6	1.2	1.2	3.8 ± 3.1	0.3	10.5	1.1	1.1	.45
T2	3.3 ± 2.4	0.3	7.5	1.4	1.4	4.4 ± 3.8	0.6	12.6	1.7	1.7	.41
All	3.3 ± 2.4	0.3	9.4	1.2	1.2	3.5 ± 2.6	0.3	9.6	1.3	1.3	.06

Note.—PD = proton density, CV = coefficient of variation.
* Data are mean ± standard deviation.



a.

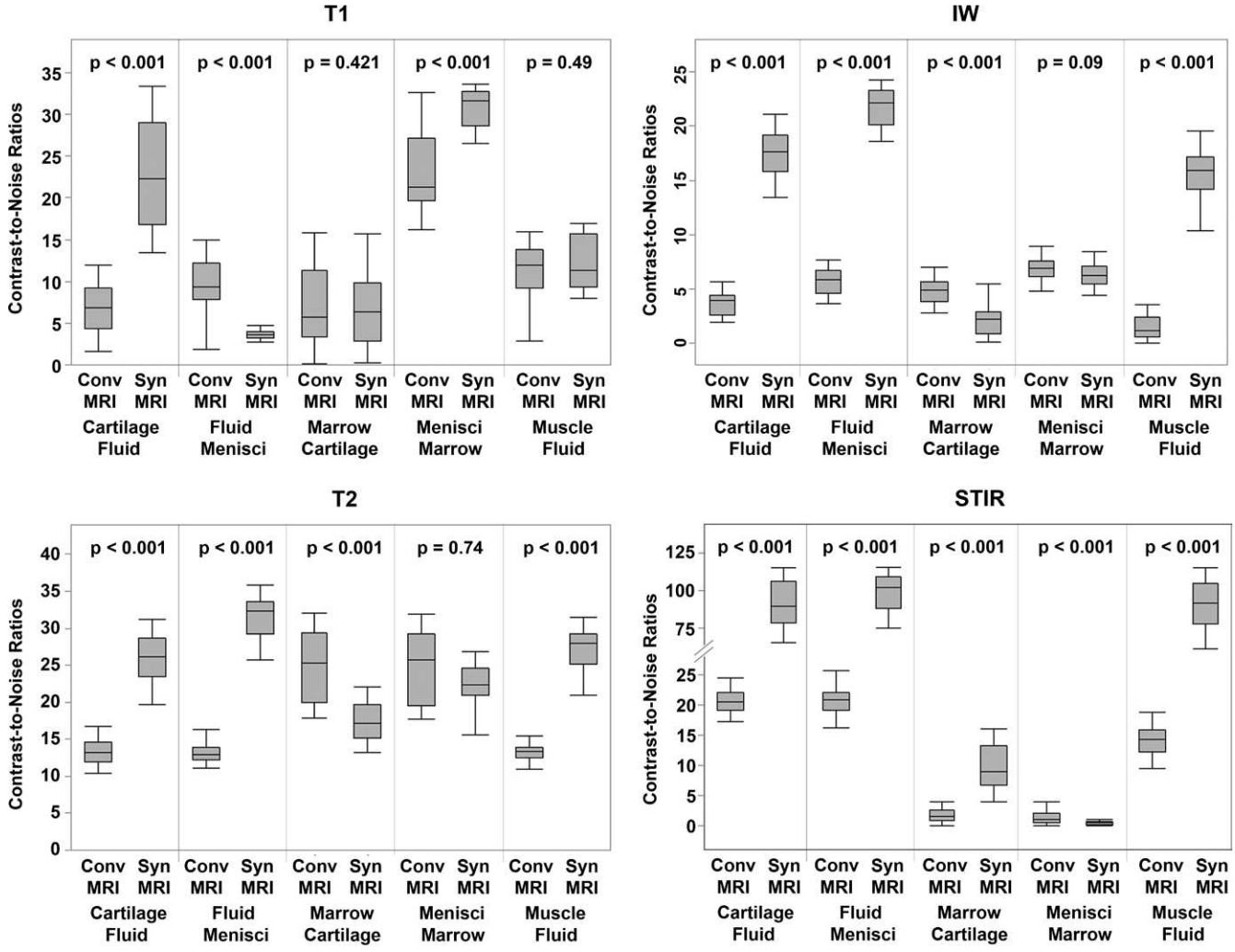
Figure 4: Box and whisker plots of **(a)** signal-to-noise ratio and **(b)** contrast-to-noise ratio of musculoskeletal tissues on synthetic and conventional morphologic MR images. *P* values refer to the comparison of a tissue type on corresponding conventional and synthetic MR images. *(Fig 4 continues)*.

Discussion

We report the native and model-corrected accuracy of synthetic knee MRI for T1, T2, and PD quantification using an ISMRM-NIST phantom and show high intraday and interday repeatability in living human participants. All synthetic MR images showed improved CNR for cartilage evaluation, and synthetic T2-weighted, intermediate-weighted, and STIR MR images showed improved CNR for meniscal evaluation. Observers perceived improvement of STIR fat suppression with synthetic MRI, whereas the overall quality ratings and detection rates of various internal knee derangements were similar with synthetic and conventional MRI.

The validation of the accuracy of synthetic MRI against a standard of reference is a prerequisite for its clinical use and appropriate patient care. Therefore, we validated and model-corrected the QRAPMASTER technique against an internationally accepted quantitative MRI phantom (13). Our approach contrasts attempts of validation that compared T2 relaxation times with other quantitative fast-spin-echo multiecho techniques (19), which introduce inaccuracies related to monoexponential T2 curve fitting (20,21) and, therefore, may not be representative of conventional single echo time fast-spin-echo T2-weighted techniques. Multiecho methods may also produce tissue-specific T2 relaxation differences when compared with conventional sequences and phantom-validated disagreements at echo times of less than 19 msec (20,22). A prior, uncalibrated phantom evaluation of a synthetic MRI prototype technique with four echo times and limited coverage of the T1, T2, and PD spectra showed T1 and T2 relaxation time underestimation and PD percentage overestimation of −1.2% ± 5.4, −6.6% ± 1.5, and 0.8% ± 1.5, respectively (23). In comparison, our phantom experiment sampled larger T1 (351–1989 msec), T2 (22–581 msec), and PD (5%–100%) domains, which better encompass the physiologic range of structures of the knee (23–25).

We show that native accuracy of the quantitative data varies in a heteroscedastic manner and that model correction should



b.

Figure 4 (continued). Conv MRI = conventional MR images, Syn MRI = synthetic MR images, IW = intermediate weighting, STIR = short-tau inversion recovery

Table 3: Observer Ratings of Image Quality Parameters of Synthetic and Conventional MRI Methods

Parameter	Conventional MRI*	Synthetic MRI*	Intermethod *P* Value	Interobserver Agreement†
Motion	5 (3, 4–5, 5)	5 (3, 4–5, 5)	.50	0.591 (0.435, 0.747)
Noise	5 (4, 4–5, 5)	4 (4, 4–5, 5)	.001	0.63 (0.483, 0.776)
Artifact	5 (4, 4–5, 5)	4 (3, 3–4, 4)	<.001	0.432 (0.282, 0.582)
Edge sharpness	4 (3, 4–5, 5)	4 (3, 4–5, 5)	.63	0.668 (0.532, 0.804)
Partial volume effects	4 (3, 3–5, 5)	4 (3, 3–5, 5)	.76	0.689 (0.573, 0.805)
Contrast resolution	4 (4, 4–5, 5)	4 (4, 4–5, 5)	.90	0.631 (0.482, 0.781)
Fluid signal	4 (4, 4–5, 5)	5 (4, 4–5, 5)	.10	0.463 (0.295, 0.63)
Fat suppression	4 (3, 3–4, 4)	5 (3, 4–5, 5)	<.001	0.57 (0.435, 0.706)

* Based on a five-point Likert scale, where 1 is the lowest value ("very bad") and 5 the highest value ("very good").

† Data are κ values, with 95% confidence intervals in parentheses.

be performed for overall reduction and error homogenization across parameter domains to maintain accuracy at the extremes of the relaxation rate curves. Such correction reduced the relative accuracy errors of measured T1, T2, and PD values from 1.9%, 7.4%, and 5.1% to 0.8%, 1.4%, and 0.3%, respectively. Our average phantom-based mean-adjusted percentage accuracy errors compare favorably with phantom-based accuracy errors of mixed-echo turbo-spin-echo techniques of 1.6%–10.9% for T1 and 9.4%–12.9% for T2 (4), 10%–13% for T2 with a multicomponent quantitative technique (26), 0%–8.3% for T1 with modified Look-Locker technique, 0%–1.2% for T1 with saturation recovery single shot acquisition

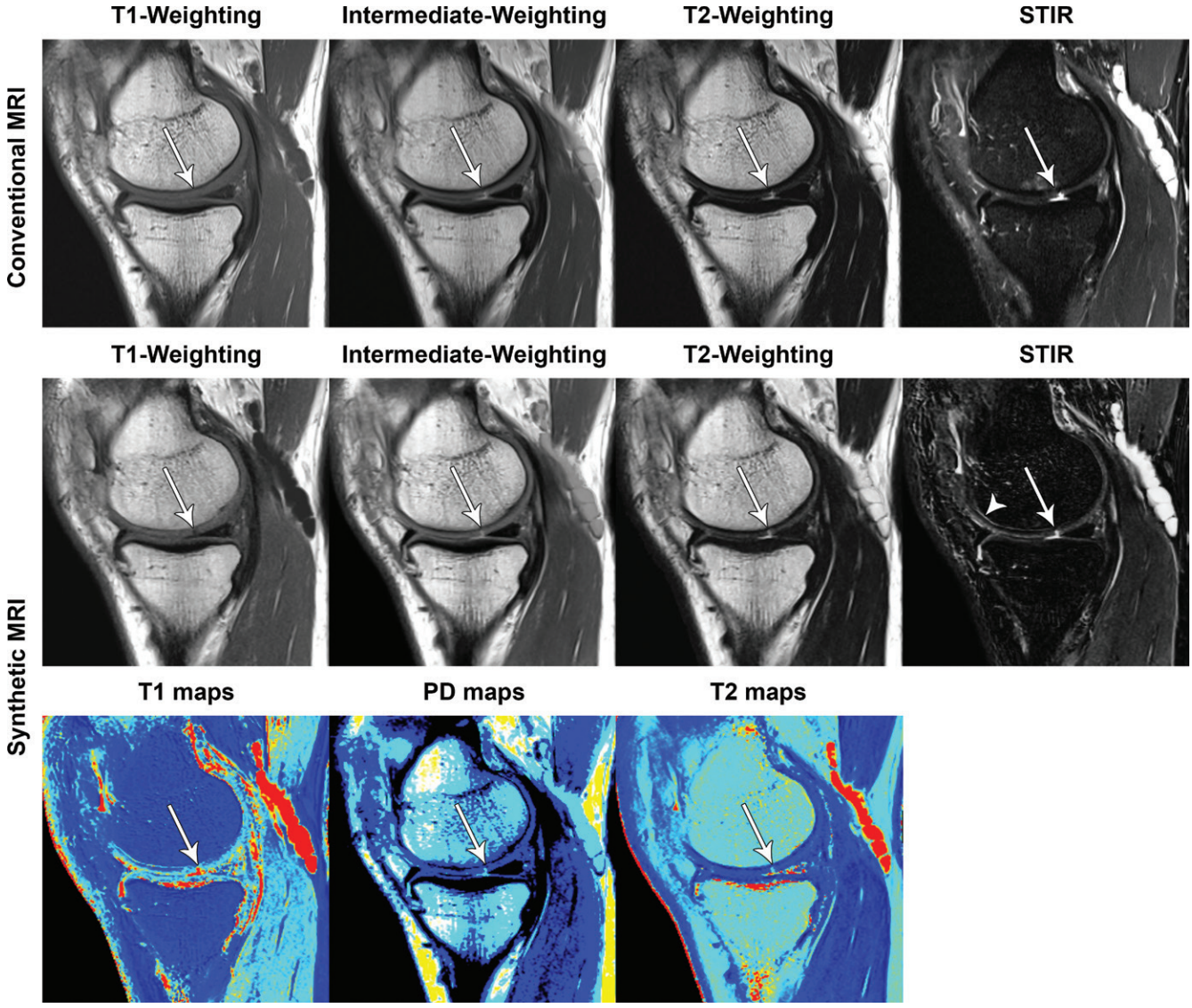


Figure 5: Images in a 51-year-old man with left knee pain. Sagittal conventional and synthetic MR images of the knee show a linear full-thickness defect of the central femoral cartilage (arrow) and a linear artifactual signal intensity along the bone-cartilage interface on the synthetic short-tau inversion recovery (STIR) image (arrowhead). PD = proton density.

technique, and 5%–15% for T2 with steady-state free precession technique (8).

We applied phantom-derived model corrections to living participants to improve in vivo accuracy. Since nonlinear model correction may unpredictably affect repeatability, we demonstrate near equivalent in vivo repeatability using split-quadratic model corrections of logarithmized data that account for heteroscedasticity and residual error structure. The phantom-based accuracy and subject-based precision errors of T1, T2, and PD quantification appear at least acceptable for clinical use. A study investigating the Osteoarthritis Initiative cohort demonstrated a significant increase of cartilage T2 relaxation times over a period of 6 years, from 32 msec to 34 msec (6.3%) in participants with simultaneous worsening in the whole organ MRI cartilage score (27). Our model-corrected T2 phantom-based accuracy error of 1.4% and subject-based precision error of 4.4% suggest the capability of our technique for detecting such a magnitude of change, which may contrast with previously reported T2 accuracy errors of 5%–15% with steady-state free precession technique (8) and 10%–13% with a Carr-Purcell-Meiboom-Grill pulse sequence (26). In addition, our corrected T1 repeatability error of 3.6% (range, 0.4%–10.1%) compares favorably to a prior conventional MRI phantom multicenter study of variable-flip-angle T1 quantification (28), which found a repeatability median error range of 0.7%–25.8% for T1 quantification.

Our initial results suggest similar detection rates with synthetic and conventional MRI for structural abnormalities of the knee; however, larger studies and correlation with arthroscopic surgery are needed to define diagnostic accuracies. Improved CNR between cartilage and fluid and menisci and fluid on synthetic T2-weighted, intermediate-weighted, and STIR MR images may help to diagnose subtle abnormalities. Synthetic MR images had a small, but higher degree of interface artifacts, which may interfere with the detection of subtle signal

Table 4: Observer Assessments of Internal Derangement with Conventional and Synthetic MRI Methods

Structural Abnormality*	Conventional MRI				Synthetic MRI				Intermethod Agreement‡		
	Frequency (n = 54)†			Interobserver Agreement‡	Frequency (n = 54)†			Interobserver Agreement‡			
	Observer 1	Observer 2	Consensus		Observer 1	Observer 2	Consensus		Observer 1	Observer 2	Consensus
Medial meniscus tear	14 (26)	16 (30)	15 (28)	0.816 (0.643, 0.989)	12 (22)	15 (28)	15 (28)	0.656 (0.422, 0.889)	0.798 (0.609, 0.987)	0.774 (0.586, 0.962)	0.723 (0.516, 0.93)
Lateral meniscus tear	13 (24)	15 (28)	14 (26)	0.807 (0.627, 0.988)	13 (24)	12 (22)	12 (22)	0.74 (0.525, 0.955)	0.797 (0.607, 0.987)	0.754 (0.551, 0.957)	0.697 (0.471, 0.922)
Articular cartilage defect	17 (31)	16 (30)	16 (30)	0.52 (0.273, 0.768)	20 (37)	17 (31)	19 (35)	0.549 (0.314, 0.784)	0.549 (0.314, 0.784)	0.695 (0.485, 0.904)	0.621 (0.397, 0.845)
Bone marrow edema	22 (41)	28 (52)	25 (46)	0.779 (0.615, 0.944)	25 (46)	24 (44)	24 (44)	0.813 (0.657, 0.969)	0.662 (0.461, 0.863)	0.705 (0.517, 0.893)	0.739 (0.558, 0.919)
All	66 (122)	75 (139)	70 (130)	0.721 (0.601, 0.841)	70 (130)	68 (126)	70 (130)	0.655 (0.523, 0.787)	0.7 (0.597, 0.803)	0.739 (0.643, 0.835)	0.704 (0.603, 0.806)

* No anterior cruciate ligament tears were seen.

† Data in parentheses are percentages.

‡ Data are κ values, with 95% confidence intervals in parentheses.

abnormalities at the tidemark of articular cartilage. We noticed improved STIR fat suppression with QRAPMASTER, which we believe is in part the result of B1 inhomogeneity correction with use of local effective flip angles (21). B1 inhomogeneity correction may also result in improved T1 contrast and account for our observation that bone marrow edema is particularly hypointense on synthetic T1-weighted MR images. As T1 hypointensity of bone marrow lesions relative to muscle is a frequently used imaging sign for marrow replacement (29), synthetic MR images may paradoxically decrease the specificity of this criterion and require additional chemical shift imaging or fat-fraction quantification for definitive evaluation. Given this finding, there is also the potential for synthetic MRI to correct for T1 bias in the fat-fraction quantification of bone marrow abnormalities without lowering flip angles, which reduces the SNR (30).

The efficiency of synthetic MRI in a clinical setting may depend on whether the total acquisition time is less than that with separately acquired conventional quantitative and morphologic MR images. In our study, synthetic and conventional MRI pulse sequence acquisition times differed by a few seconds; however, the QRAPMASTER sequence provides quantitative mapping as well as morphologic MR images in the same time that conventional MRI provides only morphologic MR images (19). While T1 mapping is most commonly used in conjunction with gadolinium-based contrast agents, T2 mapping may be the most frequently used non−gadolinium-based contrast agent technique for the detection and quantification of early cartilage degeneration. PD mapping is a promising parameter due to its association with histologic and biomechanical cartilage abnormalities (25), which can be obtained simultaneously with synthetic T2 maps. An additional potential benefit of synthetic MRI is the ability to simultaneously generate double inversion recovery images, such as STIR fluid-attenuated inversion recovery (FLAIR) images, which have been previously suggested as a replacement for postcontrast sequences in evaluating synovitis (31). However, we did not evaluate synthetic STIR FLAIR images in our study because intravenous contrast agent administration was not part of our study protocol.

Our study has limitations. We did not perform a conventional MRI comparison for the phantom experiment and did not test the derived model-correction equations in a second phantom or MR unit. Therefore, the unit- or phantom-specific systematic errors that may cause over- or undercorrection of QRAPMASTER data are unknown. However, our goal was not to produce generalizable model-correction equations, but instead to determine the in vitro accuracy improvement of the synthetic knee MRI pulse sequence with individual unit model corrections of T1, T2, and PD data and demonstrate maintained repeatability with the in vivo application of the model corrections. Additionally, our model-corrected phantom accuracies are congruent with prior synthetic and conventional MRI studies (20–22). The number of replications at each reference level, small variation associated with the replicates, and avoidance of complexity in curve fitting minimize over-fitting errors and make a training and testing set approach unnecessary. The similar proportions of internal derangement diagnosed by both

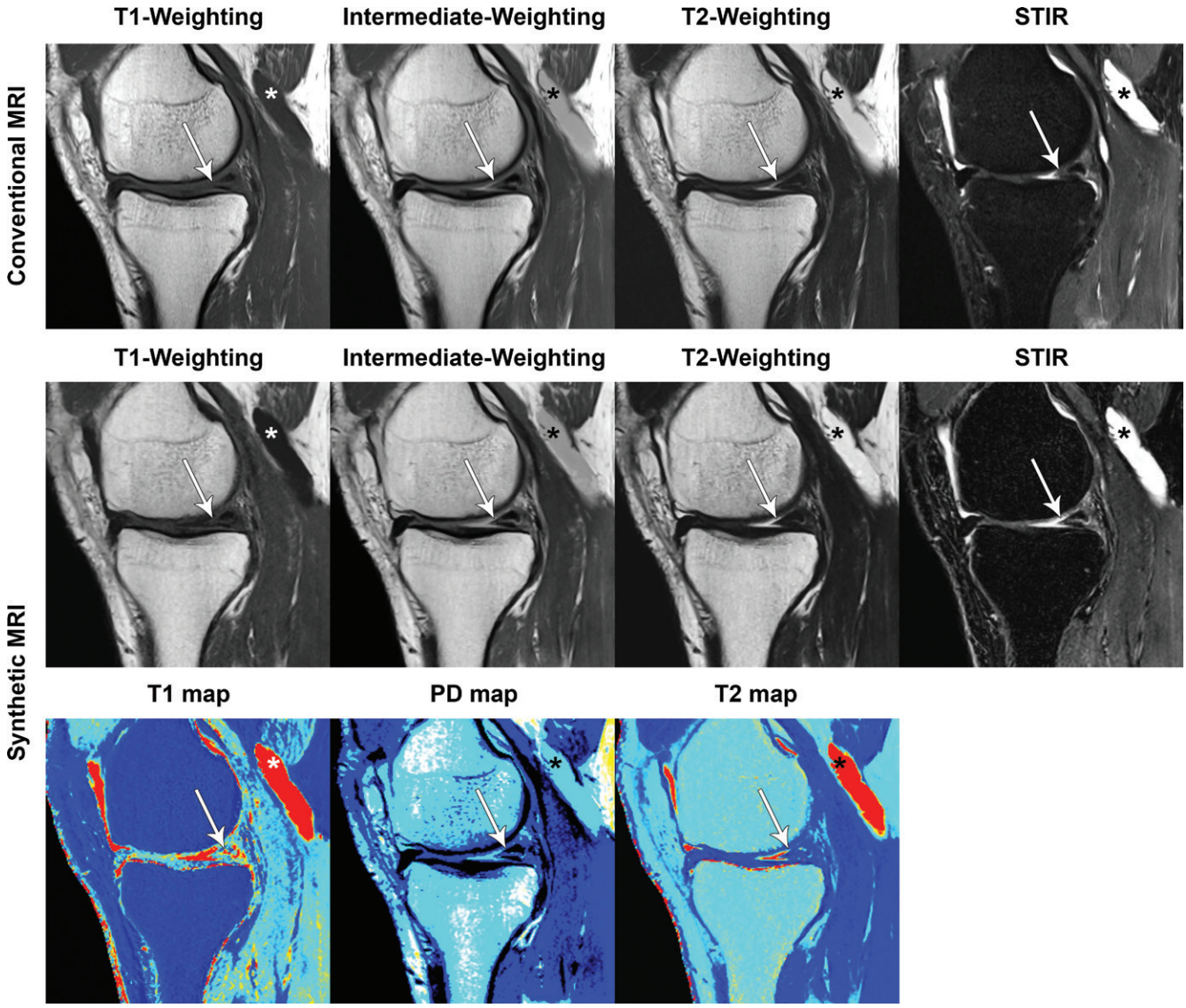


Figure 6: Images in a 44-year-old man with right knee pain. Sagittal conventional and synthetic MR images of the knee show a horizontal tear of the posterior horn of the medial meniscus (arrow). There is also a popliteal cyst (*). STIR = short-tau inversion recovery, PD = proton density.

observers with conventional and synthetic MRI suggest similar accuracies; however, agreements with surgical inspection are unknown. Owing to the size of the ISMRM-NIST phantom, we used a head coil for the phantom validation instead of the knee coil. While coil sensitivity is a static measure that T1 and T2 curve fitting compensate for (7), variations in knee position and differences of knee morphology may have contributed to lower accuracy in humans. While the heteroscedastic error calibration is a function of the measured values and therefore applicable at human body temperature, correction for residual, temperature-related, substrate-dependent errors was not possible, which may have affected the in vivo accuracy, but not repeatability and detection of structural abnormalities.

In summary, synthetic QRAPMASTER MRI of the knee is accurate for T1, T2, and PD quantification and simultaneously generates morphologic MR images with high image contrast of cartilage and meniscus relative to joint fluid and similar detection rates of structural abnormalities when compared with conventional MRI with similar acquisition time.

Acknowledgments: We thank Martin Uppman, MSc, and Tobias Granberg, MD, PhD (Karolinska University Hospital, Stockholm, Sweden), and Frederik Testud, PhD (Siemens Healthcare AB, Sweden), for their work on the QRAPMASTER pulse sequence.

Author contributions: Guarantors of integrity of entire study, N.M.K., J.B.M.W., J.F.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, N.M.K., B.F., J.B.M.W., J.F.; clinical studies, B.F., J.B.M.W., J.F.; experimental studies, B.F., J.B.M.W., Y.M.L.C., J.F.; statistical analysis, N.M.K., B.F., S.E.S., J.B.M.W., J.F.; and manuscript editing, all authors

Disclosures of Conflicts of Interest: **N.M.K.** disclosed no relevant relationships. **B.F.** disclosed no relevant relationships. **S.E.S.** disclosed no relevant relationships. **J.B.M.W.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: disclosed that he is employed part-time by and has stock in SyntheticMR AB. Other relationships: disclosed no relevant

relationships. **Y.M.L.C.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: disclosed employment with Siemens Healthcare. Other relationships: disclosed no relevant relationships. **J.F.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: disclosed grants received by institution from Siemens Healthcare and BTG International, as well as payment received by Siemens Healthcare for lectures, including service on speakers' bureaus, and travel/accommodations/meeting expenses. Other relationships: disclosed no relevant relationships.

References

1. Torrey HC. Bloch equations with diffusion terms. Phys Rev 1956;104(3):563–565.
2. Mills AF, Sakai O, Anderson SW, Jara H. Principles of quantitative MR imaging with illustrated review of applicable modular pulse diagrams. RadioGraphics 2017;37(7):2083–2105.
3. Farraher SW, Jara H, Chang KJ, Hou A, Soto JA. Liver and spleen volumetry with quantitative MR imaging and dual-space clustering segmentation. Radiology 2005;237(1):322–328.
4. Farraher SW, Jara H, Chang KJ, Ozonoff A, Soto JA. Differentiation of hepatocellular carcinoma and hepatic metastasis from cysts and hemangiomas with calculated T2 relaxation times and the T1/T2 relaxation times ratio. J Magn Reson Imaging 2006;24(6):1333–1341.
5. Kuno H, Jara H, Buch K, Qureshi MM, Chapman MN, Sakai O. Global and regional brain assessment with quantitative MR imaging in patients with prior exposure to linear gadolinium-based contrast agents. Radiology 2017;283(1):195–204.
6. Warntjes JB, Dahlqvist O, Lundberg P. Novel method for rapid, simultaneous T1, T2*, and proton density quantification. Magn Reson Med 2007;57(3):528–537.
7. Warntjes JB, Leinhard OD, West J, Lundberg P. Rapid magnetic resonance quantification on the brain: optimization for clinical usage. Magn Reson Med 2008;60(2):320–329.
8. Xanthis CG, Bidhult S, Greiser A, et al. Simulation-based quantification of native T1 and T2 of the myocardium using a modified MOLLI scheme and the importance of Magnetization Transfer. Magn Reson Imaging 2018;48:96–106.
9. Anz AW, Lucas EP, Fitzcharles EK, Surowiec RK, Millett PJ, Ho CP. MRI T2 mapping of the asymptomatic supraspinatus tendon by age and imaging plane using clinically relevant subregions. Eur J Radiol 2014;83(5):801–805.
10. Soellner ST, Goldmann A, Muelheims D, Welsch GH, Pachowsky ML. Intraoperative validation of quantitative T2 mapping in patients with articular cartilage lesions of the knee. Osteoarthritis Cartilage 2017;25(11):1841–1849.
11. Hagiwara A, Warntjes M, Hori M, et al. SyMRI of the brain: rapid quantification of relaxation rates and proton density, with synthetic MRI, automatic brain segmentation, and myelin measurement. Invest Radiol 2017;52(10):647–657.
12. Tanenbaum LN, Tsiouris AJ, Johnson AN, et al. Synthetic MRI for clinical neuroimaging: results of the Magnetic Resonance Image Compilation (MAGiC) Prospective, Multicenter, Multireader Trial. AJNR Am J Neuroradiol 2017;38(6):1103–1110.
13. Russek S, Boss M, Jackson E, et al. Characterization of NIST/ISMRM MRI System Phantom [abstr]. In: Proceedings of the Twentieth Meeting of the International Society for Magnetic Resonance in Medicine. Berkeley, Calif: International Society for Magnetic Resonance in Medicine, 2012; 2456.
14. Garnier SJ, Bilbro GL, Snyder WE, Gault JW. Noise removal from multiple MRI images. J Digit Imaging 1994;7(4):183–188.
15. Stanisz GJ, Odrobina EE, Pun J, et al. T1, T2 relaxation and magnetization transfer in tissue at 3T. Magn Reson Med 2005;54(3):507–512.
16. Tuite MJ, Kransdorf MJ, Beaman FD, et al. ACR appropriateness criteria acute trauma to the knee. J Am Coll Radiol 2015;12(11):1164–1172.
17. Eckstein F, Kunz M, Hudelmaier M, et al. Impact of coil design on the contrast-to-noise ratio, precision, and consistency of quantitative cartilage morphometry at 3 Tesla: a pilot study for the osteoarthritis initiative. Magn Reson Med 2007;57(2):448–454.
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159–174.
19. Park S, Kwack KS, Lee YJ, Gho SM, Lee HY. Initial experience with synthetic MRI of the knee at 3T: comparison with conventional T_1 weighted imaging and T_2 mapping. Br J Radiol 2017;90(1080):20170350.
20. Andreisek G, White LM, Theodoropoulos JS, et al. Synthetic-echo time postprocessing technique for generating images with variable T2-weighted contrast: diagnosis of meniscal and cartilage abnormalities of the knee. Radiology 2010;254(1):188–199.
21. Kjaer L, Thomsen C, Larsson HB, Henriksen O, Ring P. Evaluation of bi-exponential relaxation processes by magnetic resonance imaging. A phantom study. Acta Radiol 1988;29(4):473–479.
22. Hilbert T, Schulz J, Bains LJ, et al. Fast Quantitative T2 Mapping using Simultaneous-Multi-Slice and Model-Based Reconstruction [abstr]. In: Proceedings of the Twenty-Fourth Meeting of the International Society for Magnetic Resonance in Medicine. Berkeley, Calif: International Society for Magnetic Resonance in Medicine, 2016; 500.
23. Krauss W, Gunnarsson M, Andersson T, Thunberg P. Accuracy and reproducibility of a quantitative magnetic resonance imaging method for concurrent measurements of tissue relaxation times and proton density. Magn Reson Imaging 2015;33(5):584–591.
24. Gatti AA, Noseworthy MD, Stratford PW, et al. Acute changes in knee cartilage transverse relaxation time after running and bicycling. J Biomech 2017;53:171–177.
25. Nebelung S, Tingart M, Pufe T, Kuhl C, Jahr H, Truhn D. Ex vivo quantitative multiparametric MRI mapping of human meniscus degeneration. Skeletal Radiol 2016;45(12):1649–1660.
26. Maier CF, Tan SG, Hariharan H, Potter HG. T2 quantitation of articular cartilage at 1.5 T. J Magn Reson Imaging 2003;17(3):358–364.
27. Kretzschmar M, Heilmeier U, Yu A, et al. Longitudinal analysis of cartilage T2 relaxation times and joint degeneration in African American and Caucasian American women over an observation period of 6 years - data from the Osteoarthritis Initiative. Osteoarthritis Cartilage 2016;24(8):1384–1391.
28. Bane O, Hectors SJ, Wagner M, et al. Accuracy, repeatability, and interplatform reproducibility of T_1 quantification methods used for DCE-MRI: Results from a multicenter phantom study. Magn Reson Med 2018;79(5):2564–2575.
29. Richardson ML, Amparo EG, Gillespy T 3rd, Helms CA, Demas BE, Genant HK. Theoretical considerations for optimizing intensity differences between primary musculoskeletal tumors and normal tissue with spin-echo magnetic resonance imaging. Invest Radiol 1985;20(5):492–497.
30. Kühn JP, Jahn C, Hernando D, et al. T1 bias in chemical shift-encoded liver fat-fraction: role of the flip angle. J Magn Reson Imaging 2014;40(4):875–883.
31. Yoo HJ, Hong SH, Oh HY, et al. Diagnostic accuracy of a fluid-attenuated inversion-recovery sequence with fat suppression for assessment of peripatellar synovitis: preliminary results and comparison with contrast-enhanced MR imaging. Radiology 2017;283(3):769–778.